

# Experimental Test of Nonlocal Causality

M. Ringbauer<sup>1,2</sup>, C. Giarmatzi<sup>1,2</sup>, R. Chaves<sup>3,4</sup>, F. Costa<sup>1</sup>, A. G. White<sup>1,2</sup> & A. Fedrizzi<sup>1,2,5</sup>

<sup>1</sup>*Centre for Engineered Quantum Systems, <sup>2</sup>Centre for Quantum Computer and Communication Technology, School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia,*

<sup>3</sup>*Institute for Physics & FDM, University of Freiburg, 79104 Freiburg, Germany, <sup>4</sup>Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany,*

<sup>5</sup>*School of Engineering and Physical Sciences, SUPA, Heriot-Watt University, Edinburgh EH14 4AS, UK*

Explaining observations in terms of causes and effects is central to all of empirical science. Correlations between entangled quantum particles, however, seem to defy such an explanation. To recover a causal picture in this case, some of the fundamental assumptions of causal explanations have to give way. Here we consider a broad class of models where one of these assumptions, Bell’s local causality, is relaxed by allowing a direct influence from one measurement outcome to the other. We use interventional and observational data from a photonic experiment to bound the strength of this causal influence in a two-party Bell scenario and test a novel Bell-type inequality for the considered models. Our results demonstrate the incompatibility of quantum mechanics with an important class of nonlocal causal models, which includes Bell’s original model as a special case. Recovering a classical causal picture of quantum correlations thus requires an even more counter-intuitive modification of our classical notion of cause and effect.

Four decades after Freedman and Clauser [1] performed the first Bell-inequality test [2], a series of loophole-free experiments [3–5] have now conclusively shown that the predictions of quantum mechanics are at odds with the world view of local realism. Scientific realism posits that physical systems have real, objective properties— independent of whether we observe them or not—that determine the outcomes of measurements performed on the system. The idea of locality, or more precisely *local causality* is that causal influences cannot propagate faster than the speed of light. Based on local causality, and the assumption that measurement settings can be chosen freely, Bell derived an inequality that must be respected by any set of correlations that can be explained in terms of, possibly hidden, common causes, cf. Fig. 1a, but is violated by observed quantum correlations. Consequently, a new area of research has emerged, exploring to what extent the various underlying assumptions have to be relaxed in order to recover a causal explanation of quantum correlations [6–14].

An excellent platform for this research program, and a natural framework for Bell’s theorem, is the theory of causal modelling [11, 12], which aims to explain correlations in terms of cause-and-effect relations between events [15, 16]. Discovering these relations from empirical data is difficult in general [17–20], however, within classical physics such an explanation should always exist, since the properties of a classical system, even if not measured, can always be assumed to have well-defined values. Such causal reasoning is at the heart of empirical science and builds upon the most fundamental understanding of causality, that if a variable acts as the cause for another one, actively intervening on the first should cause changes in the second. More recently, causal modelling has attracted considerable interest in foundational physics, in

particular for the study of stronger-than-classical correlations [12, 21–27], dynamical causal order [28], and indefinite causal structures [28, 29] and their role as computational resource [30–33].

Phrasing Bell’s theorem in the language of causal models provides a clear picture of the underlying assumptions and allows for a unified and quantitative approach to relaxations of these assumptions [11, 12]. For example, causal models can in principle reproduce quantum correlations when relaxing Bell’s local causality assumption, which is commonly referred to as *quantum nonlocality*. Here we test models which allow for a causal influence from one measurement outcome to the other, cf. Fig. 1b. First we consider the simplest and most well-studied example of such correlations, the Clauser-Horne-Shimony-Holt (CHSH) scenario [34], where two parties, Alice and Bob, can each measure one of two dichotomic observables. Using controlled interventions we find the potential causal influence insufficiently strong to explain the observed CHSH violation. In the second experiment we go beyond the simple CHSH scenario and violate a novel Bell-type inequality which involves three measurement settings for each party and is satisfied even for arbitrarily strong causal influences from one outcome to the other [11]. In contrast to the interventional method, which requires detailed knowledge of the physical system under consideration, the latter method is device-independent. Our results highlight the incompatibility of quantum correlations, not just with Bell’s local causal model, but even with nonlocal causal models where one measurement outcome may have a direct causal influence on the other.

**Causal modeling** — A causal structure underlying  $n$  jointly distributed discrete random variables  $(X_1, \dots, X_n)$  is represented by a directed acyclic graph

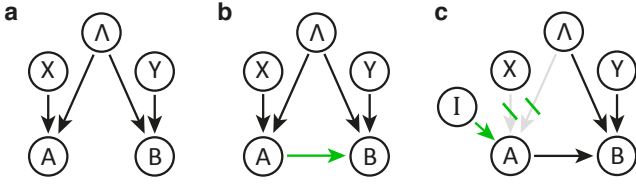


FIG. 1. **Causal structures for a Bell scenario.** (a) Bell's original local hidden variable model, where  $X$  ( $Y$ ) is Alice's (Bob's) measurement setting and  $A$  ( $B$ ) is the corresponding measurement outcome.  $\Lambda$  denotes the local hidden variable. (b) A relaxation of local causality, where  $A$  may have direct causal influence on  $B$ . Bell's model in (a) is the limiting case where the green arrow from  $A$  to  $B$  vanishes. (c) An intervention ( $I$ ) on  $A$  forces the variable to take a specific value and breaks all incoming arrows.

(DAG), with the nodes (circles in Fig. 1) representing variables and the directed edges (arrows in Fig. 1) representing causal relations [15]. Bell's theorem, where two observers, Alice and Bob, perform local measurements on one half of a shared quantum state, can be conveniently formulated in this language. Figure 1a shows the corresponding causal graph, based on Bell's assumptions of measurement independence and local causality. Measurement independence states that the measurement choices of Alice and Bob,  $X$  and  $Y$  respectively, are independent of how the system has been prepared, i.e. there is no causal link from the hidden variable  $\Lambda$  to  $X$  or  $Y$  and thus  $p(x, y, \lambda) = p(x, y)p(\lambda)$ <sup>1</sup>. Local causality implies that the probability of Alice's (Bob's) outcome  $A$  ( $B$ ) is fully specified by  $\Lambda$  and by the measurement choice  $X$  ( $Y$ ), i.e.  $p(a|x, y, b, \lambda) = p(a|x, \lambda)$  ( $p(b|y, x, a, \lambda) = p(b|y, \lambda)$ ). The latter assumption is reflected in the causal graph, Fig. 1a, by what we call *causal parameter independence*—there is no direct causal influence from the measurement setting  $Y$  ( $X$ ) to the other party's outcome  $A$  ( $B$ )—and *causal outcome independence*, stating that there is no direct causal influence from one outcome to the other<sup>2</sup>.

The causal models compatible with these assumptions are of the form of Bell's well-known local hidden variable model:  $p(a, b|x, y) = \sum_{\lambda} p(a|x, \lambda)p(b|y, \lambda)p(\lambda)$ . The constraints on the observable probabilities  $p(a, b|x, y)$  dictated by such a causal model are known as Bell inequalities. In the simplest possible Bell scenario, where each of the parties measure one of two observables ( $x, y = 0, 1$ ) obtaining one of two possible outcomes ( $a, b = 0, 1$ ), any correlations compatible with Bell's causal model must respect the CHSH-inequality [34]

$$S_2 = \langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle \leq 2, \quad (1)$$

<sup>1</sup> We adopt the usual convention that uppercase letters label random variables while their values are denoted in lower case.

<sup>2</sup> Note that these definitions differ slightly from their statistical variants, see Supplementary Information for details, see Sec. SI.

where  $\langle A_x B_y \rangle = \sum_{a,b=0,1} (-1)^{a+b} p(a, b|x, y)$  is the joint expectation value of  $A_x$  and  $B_y$ . The first loophole-free Bell experiments [3–5] now conclusively show that quantum mechanics allows for correlations that violate this inequality, therefore witnessing its incompatibility with causal models satisfying local causality and measurement independence.

In order to retain a classical causal explanation of the Bell scenario, some of these causal assumptions have to be relaxed [7–14, 35]. We focus on the class of models which do not assume causal outcome independence, such that Alice's measurement outcomes may have a direct causal influence on Bob's outcomes (or vice-versa), while satisfying causal parameter independence, see Fig. 1b. Since the causal model is formulated without any reference to a space-time structure, this influence may be sub- or superluminal, instantaneous, or even to the past, as long as it does not create any causal loop. In particular, it is consistent with a recent no-go theorem stating that quantum correlations cannot be explained by any finite-speed influence [36]. The probability distributions compatible with this causal structure can be decomposed as

$$p(a, b|x, y) = \sum_{\lambda} p(a|x, \lambda)p(b|y, a, \lambda)p(\lambda). \quad (2)$$

The first experimental method we use to test this model relies on *interventions*, a core tool in causal discovery allowing for the identification and quantification of causal influences [11, 15, 37, 38]. Formally, an intervention is the act of locally forcing a variable  $X_i$  to take on some value  $x'_i$ , denoted  $\text{do}(x'_i)$ . This removes all incoming arrows on  $X_i$ , while keeping the causal dependencies between all other variables unperturbed, see  $A$  in Fig. 1c.

In the CHSH-scenario, passive observations alone are not enough to determine whether correlations between  $A$  and  $B$  are due to direct causation or a common cause  $\Lambda$ . An intervention on variable  $A$ , however, would break the link between  $A$  and the (hypothetical) variable  $\Lambda$ . All remaining correlations between  $A$  and  $B$  must thus stem from direct causation. Indeed, the maximal shift in the probability distribution of  $B$  upon intervention on  $A$  even allows quantifying the strength of this causal link [11]. To achieve this we use the so-called *average causal effect* [15, 37],

$$\text{ACE}_{A \rightarrow B} = \sup_{b, y, a, a'} |p(b|\text{do}(a, y) - p(b|\text{do}(a'), y)|, \quad (3)$$

which is a variant of the measure  $\mathcal{C}_{A \rightarrow B}$  used in Ref. [11]. In contrast to this measure, however,  $\text{ACE}_{A \rightarrow B}$  does not require knowledge of the hidden variable and is thus experimentally accessible. As we prove in detail in Sec. SI, the average causal effect satisfies the same relation as  $\mathcal{C}_{A \rightarrow B}$  in Ref. [11], namely,

$$\min \text{ACE}_{A \rightarrow B} = \max [0, (S_2 - 2)/2], \quad (4)$$

where the maximum is taken over all eight symmetries of the CHSH quantity under relabelling of inputs, outputs,

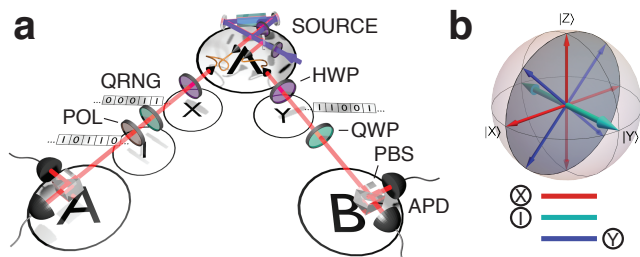


FIG. 2. **(a) The experimental setup.** Pairs of photons are generated via spontaneous parametric downconversion in a periodically poled KTP (ppKTP) crystal, using the Sagnac design of Ref. [40]. The degree of polarization entanglement between the two photons can be continuously varied by changing the polarization-angle  $\gamma$  of the pump laser. Alice and Bob perform measurements in the equatorial plane of the Bloch sphere using a half-wave plate (HWP) and a polarizing beam splitter (PBS). Additional quarter-wave plates (QWP) can be used for quantum state tomography of the initial entangled state. In the interventionist experiment an additional combination of QWP and polarizer (POL) are used between Alice's basis choice and her measurement. Causal variables are indicated using the notation of Fig. 1a. **(b)** Alice's (red) and Bob's (blue) measurement bases and the intervention direction (cyan) on the Bloch-sphere.

and parties [34]. In other words the minimal average causal effect required for a causal explanation of a set of quantum correlations is directly proportional to the CHSH violation achieved by the correlations in question.

**Interventional method** — We experimentally implemented an intervention on a CHSH-Bell test using pairs of polarization-entangled photons, generated in the state  $\cos \gamma |HV\rangle + \sin \gamma |VH\rangle$ , see Fig. 2a. Here  $H$  and  $V$  correspond to horizontal and vertical polarizations, respectively and  $\gamma$  is the polarization angle of the pump beam, which continuously controls the degree of entanglement, as measured by the concurrence  $\mathcal{C} = |\sin 2\gamma|$  [39].

Alice and Bob perform a standard CHSH-inequality test with two settings and two outcomes each. The measurements are chosen in the equatorial (linear-polarization) plane of the Bloch sphere, see Fig. 2b. In order to test the (directional) link  $A \rightarrow B$ , Bob has been located in the causal future of Alice using a 2 m fibre delay before Bob's measurement device. An intervention on Alice's outcome  $A$  can be implemented using a quarter-wave plate and a polarizer before her measurement PBS. Alice's photons are randomly projected onto circular polarization states  $|R/L\rangle = \frac{1}{\sqrt{2}}(|H\rangle \pm i|V\rangle)$ —which are orthogonal to all measurements in the performed CHSH test—and re-prepared in eigenstates of the PBS to force one of the two outcomes  $A = \pm 1$ . This approach breaks all relevant incoming causal arrows and allows deterministically setting the variable  $A$ , thus satisfying the conditions for an intervention. The measurement bases for Alice and Bob, as well as the setting of the intervention POL and QWP were chosen randomly using quantum random numbers from the Australian National Univer-

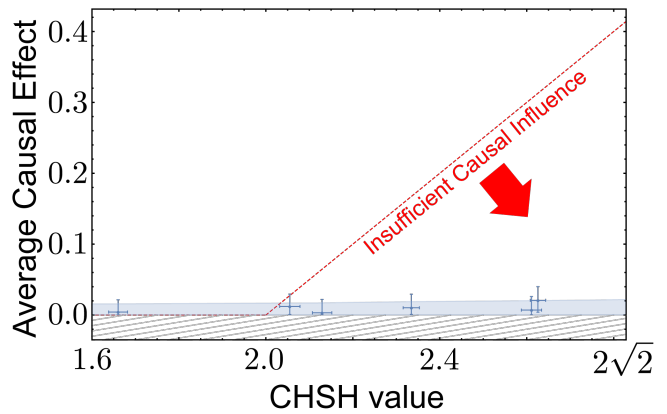


FIG. 3. **Observed average causal effect ACE versus measured CHSH-value.** Any value below the dashed red line, given by Eq. (4), is not sufficient to explain the observed CHSH-violation. Note, that the quantity ACE is bounded from below by 0, as indicated by the hatched area, resulting in asymmetric error distributions. The blue shaded area represent the  $3\sigma$  region of Poissonian noise. All errors represent  $3\sigma$  statistical confidence intervals obtained from a Monte-Carlo simulation of the Poissonian counting statistics.

sity's online quantum random number generator based on Ref. [41].

Single-photon clicks in the APDs for each outcome are registered with an AIT-TTM8000 time-tagging module with a temporal resolution of 82 ps. Outcome probabilities, used to estimate ACE, were computed from a total of 48,000 coincidence counts and no more than one event was registered for each set of random choices for  $X, Y$ , as well as the two elements of  $I$ .

Figure 3 shows the observed average causal effect as a function of the CHSH values measured for a range of entangled states. All measured values are below  $ACE_{A \rightarrow B} = 0.02^{+0.02}_{-0.02}$  and largely independent of the observed CHSH violation. Note that the quantity is bounded from below, which results in non-Gaussian statistics and makes the value 0 unachievable in the presence of experimental imperfections and finite counting statistics. When taking this in to account, all data lie within the  $3\sigma$  noise due to Poissonian counting statistics, see Sec SIV. All quoted uncertainties were obtained from Monte Carlo simulations of the Poissonian counting statistics and correspond to the 0.13<sup>th</sup> and 99.87<sup>th</sup> percentile, respectively (in the case of normal distributed variables this would correspond to  $3\sigma$  confidence regions). Within current experimental capabilities we find that CHSH violations above a value of  $S_2 = 2.05 \pm 0.02$  cannot be fully explained by means of a direct causal influence from one outcome to the other. That is, the potential causal influence between Alice's and Bob's measurement—the green arrow in Fig. 1b—is not sufficiently strong.

**Observational method** — As we have demonstrated, interventions are a powerful tool for quantifying the strength of causal influences, as measured by the av-

erage causal effect. Any experimental implementation of an intervention, however, relies on the quantum description of the degree of freedom responsible for the outcome  $A$  (in the case above, the polarization), and is thus necessarily device dependent. We now show how moving beyond the CHSH scenario allows for a device independent test of any model with an arbitrarily strong causal influence from one outcome to the other.

Consider the situation where each of the two parties can choose to measure one of *three* different dichotomic observables. As shown in Ref. [11], any correlations compatible with the model in Fig. 1b must now satisfy

$$S_3 = \langle E_{00} \rangle - \langle E_{02} \rangle - \langle E_{11} \rangle + \langle E_{12} \rangle - \langle E_{20} \rangle + \langle E_{21} \rangle \leq 4. \quad (5)$$

This inequality is symmetric and, as we show in Sec. SI, satisfied by any model that contains one-way communication of outcomes from either party to the other. Crucially, this allows us to test the models in Fig. 1b in a device-independent fashion and without committing to any particular temporal ordering of  $A$  and  $B$ .

To test inequality (5) Alice and Bob each perform measurements on their quantum system along one of three directions in the equatorial plane of the Bloch-sphere. These measurements are implemented using the setup in Fig. 2 with the intervention elements  $I$  removed. The specific measurement settings are given in Sec. SIII. Figure 4 shows the observed violation of inequality (5) as a function of the parameter  $\gamma$  of the used quantum state. The theoretical maximal violation of the inequality is achieved using a maximally entangled state, corresponding to  $\gamma = 45^\circ$ .

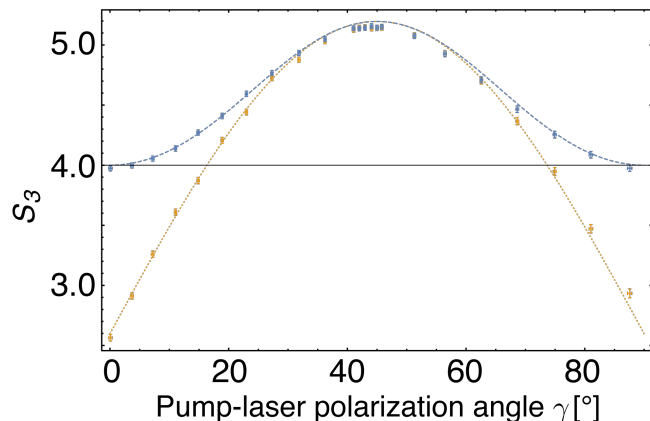


FIG. 4. **Observed values  $S_3$  for a variety of quantum states of the form  $\cos(\gamma)|HV\rangle + \sin(\gamma)|VH\rangle$ .** The orange data points are observed using a fixed measurement scheme (optimal for the maximally entangled state,  $\gamma = 45^\circ$ ), with the dotted, orange line representing the corresponding theory prediction. The blue data and blue dashed theory line corresponds to the case where measurement settings were optimized for the prepared states, see Sec. SIII. The black line represents the bound of inequality (5); any point above this line cannot be explained causally by a model of the form in Fig. 1b. Error-bars correspond to  $3\sigma$  statistical confidence intervals.

We observe a value of up to  $S_3 = 5.16_{-0.02}^{+0.02}$ , corresponding to a violation of Eq. (5) by more than 170 standard deviations. Complementary to the interventional experiment—which bounds the strength of a possible causal influence from one outcome to the other the CHSH scenario—this result rules out any causal model with an arbitrarily strong direct causal influence from one outcome to the other. This demonstrates that the conclusions of the interventional experiment hold even in a fully device-independent scenario and thus that a direct causal influence from one outcome to the other cannot explain quantum correlations.

**Discussion** — Previous work on causal explanations beyond local hidden variable models focused on testing Leggett’s *crypto nonlocality* [7, 42, 43], a notion which concerns models with a very specific choice of hidden variable and that is in fact unrelated to Bell’s local causality [44]. In contrast, we make no assumptions on the form of the hidden variable and test all models compatible with the causal structure in Fig. 1b, which is a natural generalization of Bell’s model and contains it as a special case. Practically, our experiment relies on a fair sampling assumption, see Sec. SII.

Our results demonstrate that causal modeling and interventions are powerful tools for studying causal explanations of quantum correlations beyond Bell’s local hidden variable model. It would now be of considerable interest to further develop and extend these tools to test other classes of causal models, e.g. allowing for retro-causal influences or relaxations of measurement independence [8–14, 45]. Alternatively one could completely abandon the classical notion of causality and pursue a novel framework of quantum causality [21–27]. An important question for these approaches, however, is how to treat interventions and to what extent causal discovery remains possible [46]. For example, interpretations of quantum mechanics that feature objective collapse of the wave-function would not permit interventions in the way they are used within the causal modelling framework [47].

Recent experiments put strong constraints on realist interpretations of quantum mechanics, ruling out maximally-epistemic [48] and local-causal [3–5] models. Our results exclude an important class of nonlocal causal models, thus contributing to a clearer picture of the status of reality and causality in quantum mechanics.

## ACKNOWLEDGMENTS

We thank C. Branciard, E. Cavalcanti and H. Wiseman for helpful discussions. We also thank the team from the Austrian Institute of Technology for kindly providing the time-tagging modules for this experiment. This work was supported in part by the Centres for Engineered Quantum Systems (CE110001013) and for Quantum Computation and Communication Technology (CE110001027), and the Templeton World Charity Foundation (TWCF 0064/AB38). RC acknowledges support from the Excel-

lence Initiative of the German Federal and State Governments (Grants ZUK 43 & 81), the US Army Research Office under contracts W911NF-14-1-0098 and W911NF-14-1-0133 (Quantum Characterization, Verification, and Validation), the DFG (GRO 4334 & SPP

1798). AGW acknowledges support through a UQ Vice-Chancellor's Senior Research and Teaching Fellowship, and AF through an Australian Research Council Discovery Early Career Award, DE130100240.



## Supplementary Information Experimental Test of Nonlocal Causality

Here we discuss in detail the relation of the average causal effect to the CHSH violation. We also present the derivation of our novel 3-setting inequality and discuss data and error analysis.

### SI. RELAXATION OF LOCAL CAUSALITY

Here we will discuss in detail Bell's assumption of local causality and how it relates to assumptions on the underlying causal structure.

Local causality captures the idea that there should be no causal influence from one side of the experiment to the spacelike separated other side. Formally, this is a constraint on the conditional probability distribution:  $p(a|b, x, y, \lambda) = p(a|x, \lambda)$ . Here and in the following we will not explicitly state the equivalent constraint for Bob. We would like to stress that local causality is not equivalent to signal locality, which follows from special relativity and imposes constraints on the observable probabilities only:  $p(a|x, y) = p(a|x)$ . The natural generalization of signal locality to include the hidden variable is typically referred to as *parameter independence* or *locality*:  $p(a|x, y, \lambda) = p(a|x, \lambda)$  [35]. Parameter independence together with what is often referred to *outcome independence*:  $p(a|b, x, y, \lambda) = p(a|x, y, \lambda)$  then implies local causality.

Interpreted in the causal modelling framework local causality implies that there is no causal link from Bob's measurement setting  $Y$  or outcome  $B$  to Alice's measurement outcome  $A$ , cf. Fig. S1. In the spirit of causal modelling we would like to obtain the causal structure of Bell's theorem directly from investigating these causal independencies. Specifically we denote by *causal parameter independence* the absence of a causal link from Alice's measurement setting  $X$  to Bob's measurement outcome  $B$ , and by *causal outcome independence* the absence of a causal link from Alice's measurement outcome  $A$  to Bob's measurement outcome  $B$ . As shown in Fig. S1, causal outcome independence is indeed very similar to outcome independence. In contrast to the statistical variant of parameter independence, however, causal parameter independence indeed captures the idea that there should be no direct causal influence from the measurement setting on one side to the measurement outcome on the other side. In the following we will consider in detail the causal models that satisfy causal parameter independence, but may violate causal outcome independence. The corresponding probability distribution can be decomposed as,

$$p(a, b|x, y) = \sum_{\lambda} p(a|x, \lambda) p(b|a, y, \lambda) p(\lambda). \quad (\text{S1})$$

To understand how such a model allows for the generation of nonlocal correlations, consider the two following

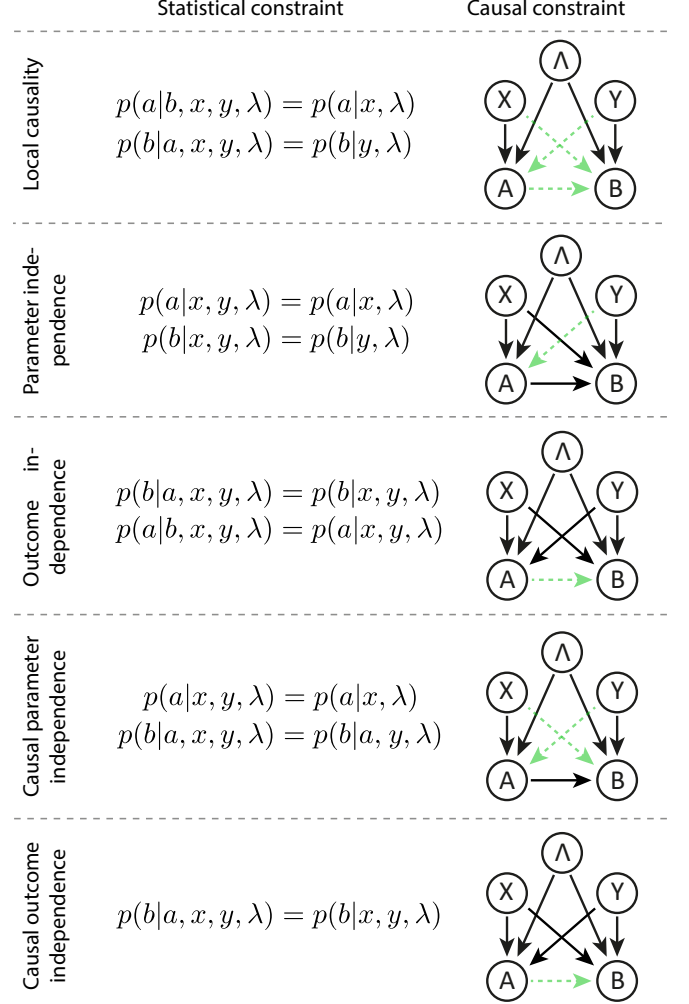


FIG. S1. **Comparison of various constraints on the causal structure of Bell's theorem.** The causal links forbidden by the respective assumption are shown in dashed green. Note that the statistical constraints implied by causal outcome independence and causal parameter independence are asymmetric in  $a$  and  $b$ , and swapping them would result in a causal structure where the arrow between  $A$  and  $B$  is reversed. Our experimental test applies to both of these structures and any convex combination of them.

deterministic strategies:

$$\begin{aligned} \text{strategy 1} &\rightarrow a = x, \quad b = a(y \oplus 1) \quad (\text{S2}) \\ \text{strategy 2} &\rightarrow a = x \oplus 1, \quad b = y(a \oplus 1) \oplus a \end{aligned}$$

Mixing the two strategies with equal probabilities generates the so-called Popescu-Rohrlich [49] distribution,  $p(a, b|x, y) = (1/2)\delta_{a \oplus b, xy}$ , which achieves maximal algebraic violation of the CHSH inequality,  $S_2 = 4$ . Crucially, however, despite the direct causal link from  $A$  to  $B$ , the above example does not permit to send signals

between Alice and Bob at the level of empirical observations, satisfying  $p(b|x, y) = p(b|x', y)$ . In this case, however, signal locality does not follow from the causal structure in Fig. 1b), but rather from the precise choice of mixing probabilities in the above strategy. This kind of *fine-tuning* of model parameters in order to ensure no-signalling has been found to be a common feature of all causal explanations of Bell correlations [12].

The concept of fine-tuning plays an important role in causal discovery, which traditionally excludes fine-tuned models as unfaithful representations. The main justification for this step stems from the fact that assuming a uniform prior over the space of probabilistic parameters, the volume of it reproducing conditional independence relations not implied by the causal structure itself has measure zero [15, 50]. In practice, however, estimation error issues due to data of finite sample size can result in the volume of unfaithful parameters being considerably large [50]. From a purely causal inference perspective, this practical aspect—together with the fact that no faithful causal model can reproduce nonlocal correlations [12]—indicates that in order to conclusively eliminate fine-tuned models as a possible causal explanation to nonlocality, new methods are required.

As in the usual Bell scenario shown in Fig. 1a), each of the probabilities appearing in Eq. (S1) can be identified with a deterministic function. To see that, consider the general case where Alice has  $m_x$  inputs and  $o_a$  outputs, that is,  $x = 0, \dots, m_x - 1$  and  $a = 0, \dots, o_a - 1$  (and analogously for Bob). The causal structure in Fig. 1b) assures that  $a = f_A(x, \lambda)$  which resembles the usual LHV model and therefore implies  $o_a^{m_x}$  different deterministic functions  $f_A$ . For  $b$ , however, we have that  $b = f_B(a, y, \lambda)$  requiring  $o_b^{o_a m_y}$  different deterministic functions  $f_B$ . That is, in order to fully describe the causal structure we need an underlying hidden variable with  $n = o_a^{m_x} o_b^{o_a m_y}$  possible values. In terms of these deterministic functions, the decomposition in (S1) can be rewritten as

$$p(a, b|x, y) = \sum_{\lambda} \delta_{a, f_A(x, \lambda)} \delta_{b, f_B(a, y, \lambda)} p(\lambda). \quad (\text{S3})$$

It is useful to represent  $p(a, b|x, y)$  as a vector  $\mathbf{p}$  with components  $\mathbf{p}_j$  labeled by the multi-index  $j = (a, b, x, y)$ . Similarly, the distribution of  $\Lambda$  can be represented by a vector with components  $\mathbf{q}_\lambda = p(\Lambda = \lambda)$ . It follows then that  $\mathbf{p} = T\mathbf{q}$  where  $T$  is a matrix with elements  $T_{j, \lambda} = \delta_{a, f_A(x, \lambda)} \delta_{b, f_B(a, y, \lambda)}$ .

#### A. Analytical derivation of the average causal effect $\min \text{ACE}_{A \rightarrow B}$ in the CHSH scenario

We now show analytically that the experimentally accessible *average causal effect* is a suitable measure of causal influence in the CHSH scenario.

In general the direct causal effect

$$\mathcal{C}_{A \rightarrow B} = \sup_{b, y, a, a'} \sum_{\lambda} p(\lambda) |p(b|do(a), y, \lambda) - p(b|do(a'), y, \lambda)|, \quad (\text{S4})$$

which quantifies the maximal shift (averaged over the unobservable variable  $\Lambda$ ) in the probability of  $B$  caused by interventions in  $A$ , can be used to quantify the strength of the causal link from  $A$  to  $B$ . Indeed, it was shown in Ref. [11] that  $\mathcal{C}_{A \rightarrow B}$  directly quantifies the degree of violation of causal outcome independence required for a causal explanation of the observed CHSH-violation as  $\min \mathcal{C}_{A \rightarrow B} = \max [0, (S_2 - 2)/2]$ , where the maximum is taken over all eight symmetries of the CHSH quantity under relabelling of inputs, outputs, and parties [34]. The quantity  $\mathcal{C}_{A \rightarrow B}$ , however, is not directly experimentally accessible, therefore precluding its use in an experimental test of the such a causal link.

Here we use an experimentally accessible variant of  $\mathcal{C}_{A \rightarrow B}$ , which does not require knowledge of the hidden variable, the *average causal effect*.

$$\begin{aligned} \text{ACE}_{A \rightarrow B} &= \sup_{b, y, a, a'} |p(b|do(a), y) - p(b|do(a'), y)| \quad (\text{S5}) \\ &= \sup_{b, y, a, a'} \pm (p(b|do(a), y) - p(b|do(a'), y)) \end{aligned}$$

This expression quantifies the average causal effect from variable  $A$  into variable  $B$ . That is, the minimum shift in the probability distribution of  $B$  that we should observe by interventions on the variable  $A$ , if indeed the underlying causal structure is that shown in Fig. 1b). We are therefore interested in the following optimization problem

$$\begin{aligned} &\underset{\mathbf{q} \in \mathbb{R}^n}{\text{minimize}} && \text{ACE}_{A \rightarrow B} \quad (\text{S6}) \\ &\text{subject to} && T\mathbf{q} = \mathbf{p} \\ & && \langle \mathbf{1}_n, \mathbf{q} \rangle = 1 \\ & && \mathbf{q} \geq \mathbf{0}_n. \end{aligned}$$

The two last constraints following from the fact that the hidden variable  $\Lambda$  should be described by a well defined probability distribution (positive and normalized).

To write the optimization problem Eq.(S6) as a standard linear program, notice that

$$(p(b|do(a), y) - p(b|do(a'), y)) \quad (\text{S7})$$

$$= \sum_{\lambda} p(\lambda) (\delta_{b, f_B(a, y, \lambda)} - \delta_{b, f_B(a', y, \lambda)}) \quad (\text{S8})$$

$$= \sum_i q_i v_i = \langle \mathbf{v}, \mathbf{q} \rangle, \quad (\text{S9})$$

where the vector  $\mathbf{v} = \mathbf{v}(a, a', y, b)$  fully characterizes the action of the Kronecker-symbols in eq. (S8). The ACE measure (S5) can then be recast as

$$\text{ACE}_{A \rightarrow B} = \max_{i=1, \dots, 2L} \langle \mathbf{q}, \mathbf{v}_i \rangle = C\mathbf{q}. \quad (\text{S10})$$

Here, the index  $i$  parametrizes the  $2L$  possible instances of  $(a, a', y, b)$  with  $x \neq x'$  (the factor 2 coming from

the  $\pm$  signs in Eq. (S5)) and  $\mathbf{v}_i = v(a, a', y, b)$  denotes the vector corresponding to that instance. The matrix  $C$  subsumes all these different instances, that is,  $C := \sum_{i=1}^L |\mathbf{e}_i\rangle\langle\mathbf{v}_i|$  where  $\mathbf{e}_i$  stands for an orthonormal basis.

The optimization problem in Eq. (S6) is then equivalent to a standard linear program

$$\begin{aligned} & \underset{\mathbf{v}, \mathbf{q}}{\text{minimize}} && v \\ & \text{subject to} && T\mathbf{q} = \mathbf{p} \\ & && \langle \mathbf{1}_n, \mathbf{q} \rangle = 1 \\ & && \mathbf{q} \geq \mathbf{0}_n \\ & && C\mathbf{q} \leq v. \end{aligned} \quad (\text{S11})$$

As proved in Ref. [11], to obtain the solution of this problem for any vector  $\mathbf{p}$  encoding the full probability distribution  $p(a, b|x, y)$ , we have to consider the dual optimization problem. In practice, that means that solving Eq. (S11) is equivalent to evaluating

$$\max_{1 \leq i \leq K} \langle \mathbf{d}_i, \mathbf{p} \rangle,$$

where  $\{\mathbf{d}_i\}_{i=1}^K$  denotes the vertices of the dual feasible region (see Ref. [11] for further details).

We have performed such an analysis for the particular case of the CHSH scenario ( $m_x = m_y = o_a = o_b = 2$ ) using PORTA [51], a standard software for the evaluation of extremal points of a polyhedron. Similarly to what happens to the measure  $\mathcal{C}_{A \rightarrow B}$  (see Supplemental material of Ref. [11]), the extremal points of the dual region correspond to non-signalling constraints and all the symmetries of the CHSH inequality (up to a constant factor), implying that

$$\min \text{ACE}_{A \rightarrow B} = \max [0, (S - 2)/2]. \quad (\text{S12})$$

### B. Proving the new inequality

As shown above, a relaxation of causal outcome independence, while maintaining causal parameter independence, allows for the classical explanation of any nonlocal correlations in the CHSH scenario, where the two parties perform two possible dichotomic measurements. However, if the parties perform three or more measurements each, this does not hold anymore.

The decomposition (S3) defines a polytope of correlations that are compatible with the causal model in Fig. 1b. This polytope is characterized by  $n = o_a^{m_x} o_b^{o_a m_y}$  extremal points. Therefore, for a fixed number of measurements and outcomes, one can resort to usual convex optimization software in order to find its description in terms of finitely many Bell inequalities. As shown in Ref. [11], one of the Bell inequalities characterizing the polytope in the case  $o_a = o_b = 2$  and  $m_x = m_y = 3$  is given by inequality (5). An easy way of proving that this inequality in fact holds, is to verify that for each of

the  $n = 2^3 2^6$  extremal points defining the polytope this inequality is satisfied.

A similar argument can be used to prove that this inequality is also valid if we reverse the roles of Alice and Bob. That is, in this case we allow the outcomes of Alice to depend on the outcomes of Bob:

$$p(a, b|x, y) = \sum_{\lambda} p(a|x, b, \lambda) p(b|y, \lambda) p(\lambda). \quad (\text{S13})$$

Since both the models Eq. (S1) and Eq. (S13) respect inequality (5), so does a convex combination of both of them. In other words, any model of the form

$$p(a, b|x, y) = \sum_{\lambda} p(a|x, \lambda) p(b|y, a, \lambda) p(\lambda) + \sum_{\mu} p(a|x, b, \mu) p(b|y, \mu) p(\mu), \quad (\text{S14})$$

with  $\sum_{\mu, \lambda} p(\mu) + p(\lambda) = 1$ , respects inequality (5). Any pure two-qubit entangled state, however, can generate correlations violating this inequality [11]. This allows us to show unambiguously and based on observational data only, that a direct causal influence from one outcome to the other cannot explain quantum correlations.

## SII. THEORETICAL ANALYSIS OF EXPERIMENTAL IMPERFECTIONS

Any experiment suffers from imperfections in the form of detector inefficiencies, noise and other possible forms of loss. Practically our experiment thus relies on a fair-sampling assumption. Here we provide a short analysis of the requirements for testing inequality (5) without this assumption. Our analysis is similar to what is usually done for Bell inequalities, for example, the CHSH inequality [52, 53]

To describe the inefficiency of the photon detectors, we model the projective measurements being performed in the experiment via the following POVM with elements [54]

$$\begin{aligned} M_{\uparrow} &= \eta_{\uparrow} |\uparrow\rangle_s \langle\uparrow| + (1 - \eta_{\downarrow}) |\downarrow\rangle_s \langle\downarrow|, \\ M_{\downarrow} &= \eta_{\downarrow} |\downarrow\rangle_s \langle\downarrow| + (1 - \eta_{\uparrow}) |\uparrow\rangle_s \langle\uparrow|. \end{aligned} \quad (\text{S15})$$

For perfect efficiencies  $\eta_{\downarrow, \uparrow} = 1$  this POVM implements a projective measurement along the direction defined by  $s$ , that is,  $|\uparrow\rangle_s, |\downarrow\rangle_s$  are the eigenstates of the observable  $|\uparrow\rangle_s \langle\uparrow| - |\downarrow\rangle_s \langle\downarrow|$ . Following the usual treatment of detection inefficiencies [52–54], given a non-click event of the detectors happening with probability  $\eta$  (assumed to be the same for both detectors), we bin it together with the  $\downarrow$ -outcome. That is, our faulty measurements are described by Eq. (S15) with  $\eta_{\downarrow} = 1$  and  $\eta_{\uparrow} = \eta$ .

We also consider errors in the preparation of the states. To that aim we consider the initial two-qubit state to be



affected by white noise (parameterized by the visibility  $v$ ):

$$\rho_\gamma(v) = v|\Psi_\gamma\rangle\langle\Psi_\gamma| + (1-v)\mathbb{I}/4 \quad (\text{S16})$$

where  $\Psi_\gamma = \cos\gamma|HV\rangle + \sin\gamma|VH\rangle$  represents the pure two-qubit entangled state that we would ideally prepare.

For different values of  $\gamma$  we have performed a numerical optimization over all possible POVMs Eq. (S15) in order to find the minimum values of  $v$  and  $\eta$  leading to the violation of inequality (5) without requiring the fair-sampling assumption. The results are shown in Fig. S2. As can be seen, the smaller is the initial entanglement (parameterized by the angle  $\theta$ ) the less robust is the violation of the inequality as function of the visibility and detector inefficiency. For a maximally entangled state the minimum required visibility and efficiency are, respectively,  $v_{\text{crit}} \sim 0.77$  and  $\eta_{\text{crit}} \sim 0.88$ . For comparison, we also compute the requirements for the usual CHSH inequality [52, 53]. Not surprisingly, since our inequality excludes a larger class of causal models, it has more stringent experimental requirements.

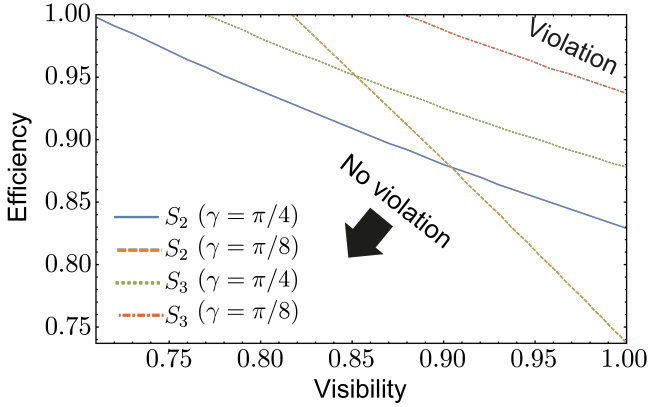


FIG. S2. **Efficiency  $\eta$  and visibility  $v$  requirements for a violation of inequality (5) without fair-sampling assumption.** For a maximally entangled state ( $\gamma = \pi/4$ ) inequality (5) is violated above the green, dashed line, while the CHSH inequality is violated above the blue, solid line. For a non-maximally entangled state ( $\gamma = \pi/8$ ) the corresponding lines are red dot-dashed for inequality (5) and orange, dashed for CHSH.

### SIII. TESTING THE NEW INEQUALITY

Similarly to the standard CHSH inequality, our 3-setting inequality (5) can be tested using only measurements in the equatorial plane of the form  $O = \cos(\theta)\hat{Z} + \sin(\theta)\hat{X}$ . Numerical results suggest that this is indeed optimal and the maximal violation of the inequality is achieved for  $\theta_0^{(A)} = -\pi/6$ ,  $\theta_1^{(A)} = 7\pi/6$ ,  $\theta_2^{(A)} = \pi/2$ ,  $\theta_0^{(B)} = -\pi/3$ ,  $\theta_1^{(B)} = \pi/3$  and  $\theta_2^{(B)} = \pi$ , where the superscript denotes the party and the subscript denotes the number of the

measurement. Using these measurement settings we obtain

$$S_3 = \frac{3}{2}\sqrt{3}(1 + \sin(2\gamma)).$$

Note that since  $\mathcal{C} = |\sin 2\gamma|$  this corresponds to a linear relationship between  $S_3$  and the concurrence of the used state.

As pointed out earlier, inequality (5) can indeed be violated by every entangled quantum state. This is clearly demonstrated by  $\theta_0^{(A)} = -\alpha$ ,  $\theta_1^{(A)} = \alpha + \pi$ ,  $\theta_2^{(A)} = \frac{\pi}{2}$ ,  $\theta_0^{(B)} = -\beta$ ,  $\theta_1^{(B)} = \beta$  and  $\theta_2^{(B)} = \pi$ , where the optimal angles  $\alpha, \beta$  can be found analytically as a function of the state parameter  $\gamma$ . In this case we obtain:

$$S_3 = \left( \frac{\sqrt{\cos(4\gamma) + 7} - \sqrt{2}(\cos(2\gamma) - 3)}{2\cos^2(\gamma)} \right) \times \left( \sqrt{\cos(2\gamma) + \sqrt{2}\sqrt{\cos(4\gamma) + 7} - 3} \right)$$

The corresponding angles  $\alpha$  and  $\beta$  are shown in Fig. S3.

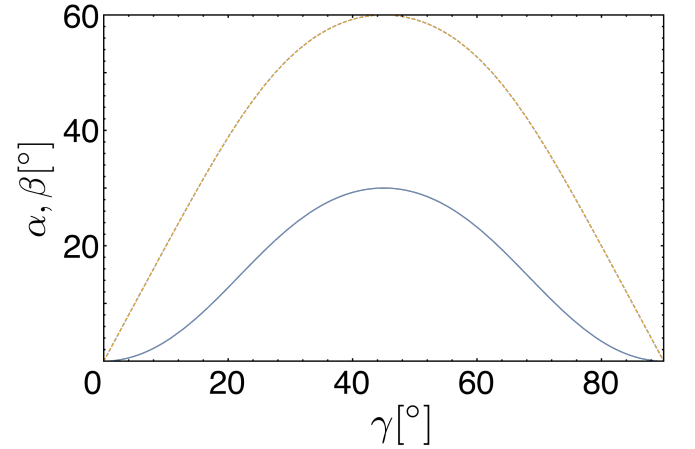


FIG. S3. **Measurement angles for inequality (5)** Alice's measurement angle  $\alpha$  shown in solid, blue, and Bob's measurement angle  $\beta$  in dashed, orange.

### SIV. ERROR ANALYSIS

Here we discuss the details of the error analysis for both experiments above.

Any photonic experiment suffers from statistical noise due to the Poissonian nature of the single-photon source and detection. Additionally it is a common feature for experiments measuring bounded quantities (such as ACE) that observed distributions feature significant asymmetry close to the boundaries, making the observation of extreme values very unlikely. To illustrate this effect, Fig. S4 shows simulated statistics for ACE assuming perfect measurement and only Poissonian noise at a total

number of 48000 single-photon events. For our choice of quantum states, the distribution of single-photon clicks is such, that higher entangled states are more susceptible to this form of noise. As a consequence, the median of the distribution increases with entanglement from 0.0048 to 0.0068, with  $3\sigma$  confidence intervals of [0.0002, 0.0155] and [0.0003, 0.0219], respectively. Note that all  $3\sigma$ -intervals quoted in this manuscript correspond to the intervals that contain  $\sim 99.73\%$  of the data (in analogy with the  $3\sigma$ -region for a normal distribution) and are thus asymmetric around the median of the distribution, see Fig. S4.

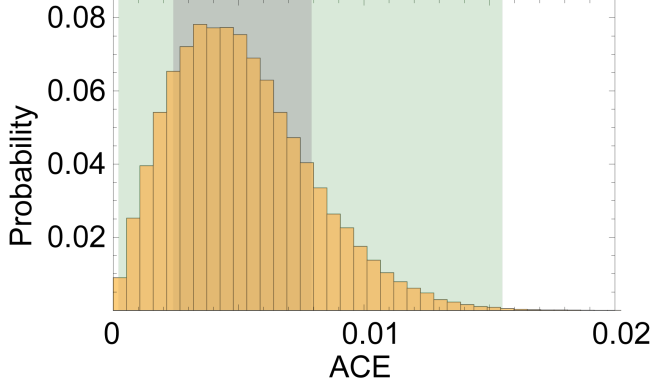


FIG. S4. **Distribution of statistical noise due to Poissonian counting statistics.** The dark, purple shaded area corresponds to the 68.27% ( $1\sigma$ ) confidence interval, while the light, green area is the 99.73% ( $3\sigma$ ) interval. These have been chosen for comparability with normal distributed data. Data was obtained from 100,000 runs of a Monte-Carlo simulation of the Poissonian counting statistics for perfect measurements.

Besides the statistical errors there are various sources of systematic errors, which explain the consistent offset

of the measured ACE from 0. For the interventionist experiment it is crucial that all CHSH measurements are performed in the equatorial plane, while the intervention acts on the poles of the Bloch sphere. The relevant waveplates are one HWP each to set Alice' and Bob's measurement basis, and one QWP for the intervention. The relative phase-shifts imparted by these waveplates and their accuracy are listed in Tab. I. These errors result in a tilt of the intervention from the orthogonal orientation of  $0.0109^{+0.0122}_{-0.0009}$  and Alice's and Bob's measurement planes are tilted by  $0.1403^{+0.0004}_{-0.0005}$  and  $0.050^{+0.013}_{-0.010}$ , respectively.

The accuracy of the intervention polarizer is  $\Delta\phi = 0.14^\circ$ , with a contrast better than 7000:1. Alice's and Bob's measurement PBS have a contrast of greater than 500 : 1 and 8000 : 1, respectively and are aligned to each other to a contrast of 1000 : 1. All of these systematic and statistical errors were taken into account in a Monte Carlo simulation. We observe only a slight increase in the median values and variance of the expected uncertainty distributions, and the systematic imperfections result in an additional offset from 0. This suggests that the offset

Element	$\Delta\phi(^{\circ})$	$r$
HWP (X)	0.04	0.955
HWP (Y)	0.04	0.978
QWP (I)	0.13	0.503

TABLE I. Shown are the standard deviations  $\Delta\phi$  in the angles of the optical axes as determined from fits to the measured coincidence counts, as well as the retardance  $r$  (in measures of  $\lambda/2$ ) obtained from the visibility of the fringes observed in the same data, for the relevant waveplates.

in the data of Fig. 3 could be explained by systematic imperfections.

- 
- [1] S. J. Freedman and J. F. Clauser, Phys. Rev. Lett. **28**, 938 (1972).
  - [2] J. S. Bell, Physics **1**, 195 (1964).
  - [3] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenber, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, et al., Nature **526**, 682 (2015).
  - [4] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, et al., Phys. Rev. Lett. **115**, 250402 (2015).
  - [5] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán, et al., Phys. Rev. Lett. **115**, 250401 (2015).
  - [6] C. H. Brans, Int. J. Theor. Phys. **27**, 219 (1988).
  - [7] C. Branciard, N. Brunner, N. Gisin, C. Kurtsiefer, A. Lamas-Linares, A. Ling, and V. Scarani, Nat. Phys. **4**, 681 (2008).
  - [8] M. J. W. Hall, Phys. Rev. Lett. **105**, 250404 (2010).
  - [9] M. J. W. Hall, Phys. Rev. A **84**, 022102 (2011).
  - [10] J. Barrett and N. Gisin, Phys. Rev. Lett. **106**, 100406 (2011).
  - [11] R. Chaves, R. Kueng, J. B. Brask, and D. Gross, Phys. Rev. Lett. **114**, 140403 (2015).
  - [12] C. J. Wood and R. W. Spekkens, New J. Phys. **17**, 033002 (2015).
  - [13] D. Aktas, S. Tanzilli, A. Martin, G. Pütz, R. Thew, and N. Gisin, Phys. Rev. Lett. **114**, 220404 (2015).
  - [14] G. Pütz and N. Gisin, *Preprint at <http://arxiv.org/abs/1510.09087>* (2015).
  - [15] J. Pearl, *Causality*. (Cambridge University Press, 2009).
  - [16] P. Spirtes, N. Glymour, and R. Scheines, *Causation, Prediction, and Search*, 2nd ed. (The MIT Press, 2001).
  - [17] D. Geiger and C. Meek, in *Proceedings of the 15th conference on Uncertainty in Artificial Intelligence* (1999), pp. 226–235.
  - [18] J. Tian and J. Pearl, in *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence* (Morgan Kaufmann Publishers Inc., 2002), pp. 519–527.

- [19] R. Chaves, L. Luft, T. O. Maciel, D. Gross, D. Janzing, and B. Schölkopf, Proceedings of the 30th Conference on Uncertainty in Artificial Intelligence pp. 112–121 (2014).
- [20] J. M. Mooij, J. Peters, D. Janzing, J. Zscheischler, and B. Schölkopf, *Preprint at <http://arxiv.org/abs/1412.3773>* (2014).
- [21] R. Chaves, C. Majenz, and D. Gross, *Nat. Commun.* **6**, 5766 (2015).
- [22] T. Fritz, *New J. Phys.* **14**, 103001 (2012).
- [23] T. Fritz, *Comm. Math. Phys.* pp. 1–45 (2015).
- [24] E. G. Cavalcanti and R. Lal, *J. Phys. A: Math. Theor.* **47**, 424018 (2014).
- [25] J. Pienaar and Č. Brukner, *New J. Phys.* **17**, 073020 (2015).
- [26] M. S. Leifer and R. W. Spekkens, *Physical Review A* **88**, 052130 (2013).
- [27] J. Henson, R. Lal, and M. F. Pusey, *New J. Phys.* **16**, 113043 (2014).
- [28] O. Oreshkov and C. Giarmatzi, *Preprint at <http://arxiv.org/abs/1506.05449v2>* (2015).
- [29] Č. Brukner, *Nat. Phys.* **10**, 259 (2014).
- [30] G. Chiribella, G. M. D’Ariano, P. Perinotti, and B. Valiron, *Phys. Rev. A* **88**, 022318 (2013).
- [31] G. Chiribella, *Phys. Rev. A* **86**, 040301 (2012).
- [32] M. Araújo, F. Costa, and Č. Brukner, *Phys. Rev. Lett.* **113**, 250402 (2014).
- [33] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. a. Rozema, Č. Brukner, and P. Walther, *Nat. Commun.* **6**, 7913 (2015).
- [34] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Phys. Rev. Lett.* **23**, 880 (1969).
- [35] H. M. Wiseman and E. G. Cavalcanti, *Preprint at <http://arxiv.org/abs/1503.06413>* (2015), to be published in *Quantum (Un)speakables II*.
- [36] J.-D. Bancal, S. Pironio, A. Acín, Y.-C. Liang, V. Scarani, and N. Gisin, *Nat. Phys.* **8**, 867 (2012).
- [37] D. Janzing, D. Balduzzi, M. Grosse-Wentrup, and B. Schölkopf, *Ann. Statist.* **41**, 2324 (2013).
- [38] K. Ried, M. Agnew, L. Vermeyden, D. Janzing, R. W. Spekkens, and K. J. Resch, *Nat. Phys.* **11**, 414 (2015).
- [39] S. Hill and W. K. Wootters, *Phys. Rev. Lett.* **78** (1997).
- [40] A. Fedrizzi, T. Herbst, A. Poppe, and A. Zeilinger, *Opt. Exp.* **15**, 15377 (2007).
- [41] T. Symul, S. M. Assad, and P. K. Lam, *App. Phys. Lett.* **98** (2011).
- [42] S. Gröblacher, T. Paterek, R. Kaltenbaek, Č. Brukner, M. Zukowski, M. Aspelmeyer, and A. Zeilinger, *Nature* **446**, 871 (2007).
- [43] T. Paterek, A. Fedrizzi, S. Gröblacher, T. Jennewein, M. Zukowski, M. Aspelmeyer, and A. Zeilinger, *Phys. Rev. Lett.* **99**, 210406 (2007).
- [44] C. Branciard, *Phys. Rev. A* **88**, 042113 (2013).
- [45] J. Gallicchio, A. S. Friedman, and D. I. Kaiser, *Phys. Rev. Lett.* **112**, 110405 (2014).
- [46] F. Costa and S. Shrapnel, *Preprint at <http://arxiv.org/abs/1512.07106>* (2015).
- [47] P. M. Näger, *Synthese* pp. 1–29 (2015).
- [48] M. Ringbauer, B. Duffus, C. Branciard, E. G. Cavalcanti, A. G. White, and A. Fedrizzi, *Nature Physics* **11**, 249 (2015).
- [49] S. Popescu and D. Rohrlich, *Found. Phys.* **24**, 379 (1994).
- [50] C. Uhler, G. Raskutti, P. Bühlmann, B. Yu, et al., *Ann. Statist.* **41**, 436 (2013).
- [51] T. Christof and A. Löbel, *PORTA – Polyhedron Representation Transformation Algorithm* (2009).
- [52] P. H. Eberhard, *Phys. Rev. A* **47**, R747 (1993).
- [53] J. Wilms, Y. Disser, G. Alber, and I. C. Percival, *Phys. Rev. A* **78**, 032116 (2008).
- [54] R. Chaves and J. B. Brask, *Phys. Rev. A* **84**, 062110 (2011).